Unary profile of lambda terms with restricted De Bruijn indices

K. Grygiel[1] I. Larcher[1]

[1] Institute of Discrete Mathematics and Geometry
Technische Universität Wien

CLA Versailles, 2019

Lambda terms

Let $\mathcal{V}$ be a countable set of variables. Lambda terms are defined by the following grammar:

$$\mathcal{T} ::= \mathcal{V} \mid \lambda\mathcal{V}.\mathcal{T} \mid \mathcal{T}\ \mathcal{T}$$

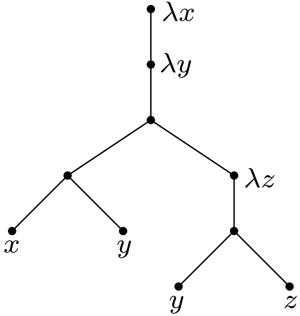
Lambda terms

Let $\mathcal{V}$ be a countable set of variables. Lambda terms are defined by the following grammar:

$$\mathcal{T} ::= \mathcal{V} \mid \lambda\mathcal{V}.\mathcal{T} \mid \mathcal{T}\ \mathcal{T}$$

Example:

$\lambda x.\lambda y.((xy)(\lambda z.(yz)))$



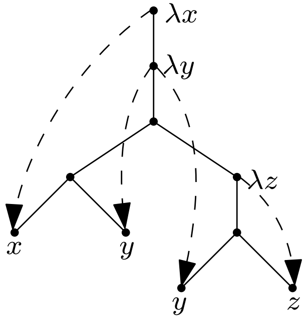
Lambda terms

Let $\mathcal{V}$ be a countable set of variables. Lambda terms are defined by the following grammar:

$$\mathcal{T} ::= \mathcal{V} \mid \lambda\mathcal{V}.\mathcal{T} \mid \mathcal{T}\ \mathcal{T}$$

Example:

$\lambda x.\lambda y.((xy)(\lambda z.(yz)))$



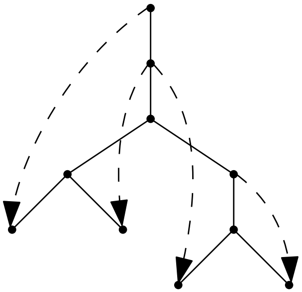
Lambda terms

Let $\mathcal{V}$ be a countable set of variables. Lambda terms are defined by the following grammar:

$$\mathcal{T} ::= \mathcal{V} \mid \lambda\mathcal{V}.\mathcal{T} \mid \mathcal{T}\ \mathcal{T}$$

Example:

$\lambda x.\lambda y.((xy)(\lambda z.(yz)))$



Lambda terms

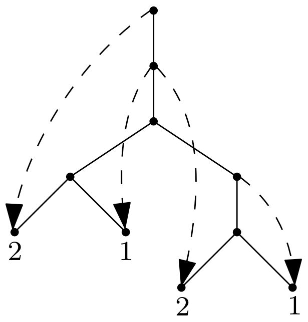
Let $\mathcal{V}$ be a countable set of variables. Lambda terms are defined by the following grammar:

$$\mathcal{T} ::= \mathcal{V} \mid \lambda\mathcal{V}.\mathcal{T} \mid \mathcal{T}\ \mathcal{T}$$

Example:

$\lambda x.\lambda y.((xy)(\lambda z.(yz)))$

De Bruijn notation:
$\lambda\lambda((21)(\lambda(21)))$



Lambda terms

Let $\mathcal{V}$ be a countable set of variables. Lambda terms are defined by the following grammar:

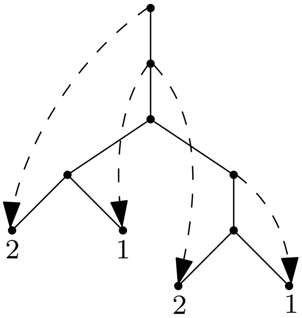
$$\mathcal{T} ::= \mathcal{V} \mid \lambda\mathcal{V}.\mathcal{T} \mid \mathcal{T}\ \mathcal{T}$$

Example:

$\lambda x.\lambda y.((xy)(\lambda z.(yz)))$

De Bruijn notation:
$\lambda\lambda((21)(\lambda(21)))$

1,2 ... De Bruijn indices



Lambda terms

Let $\mathcal{V}$ be a countable set of variables. Lambda terms are defined by the following grammar:

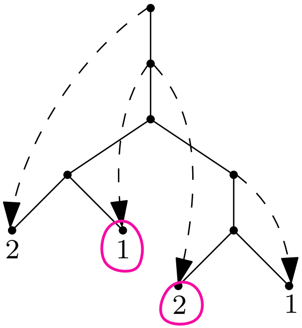
$$\mathcal{T} ::= \mathcal{V} \mid \lambda\mathcal{V}.\mathcal{T} \mid \mathcal{T}\ \mathcal{T}$$

Example:

$\lambda x.\lambda y.((xy)(\lambda z.(yz)))$

De Bruijn notation:
$\lambda\lambda((21)(\lambda(21)))$

1,2 ... De Bruijn indices



Lambda terms

Let $\mathcal{V}$ be a countable set of variables. Lambda terms are defined by the following grammar:

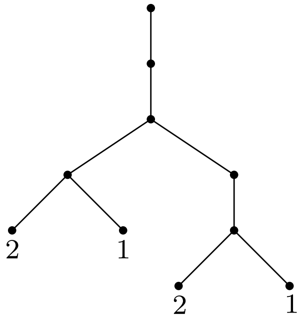
$$\mathcal{T} ::= \mathcal{V} \mid \lambda\mathcal{V}.\mathcal{T} \mid \mathcal{T}\ \mathcal{T}$$

Example:

$\lambda x.\lambda y.((xy)(\lambda z.(yz)))$

De Bruijn notation:
$\lambda\lambda((21)(\lambda(21)))$

1,2 . . . De Bruijn indices



Lambda terms

Let $\mathcal{V}$ be a countable set of variables. Lambda terms are defined by the following grammar:

$$\mathcal{T} ::= \mathcal{V} \mid \lambda\mathcal{V}.\mathcal{T} \mid \mathcal{T}\ \mathcal{T}$$

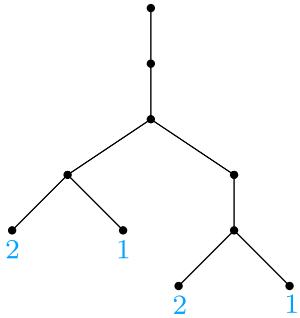
Example:

$\lambda x.\lambda y.((xy)(\lambda z.(yz)))$

De Bruijn notation:
$\lambda\lambda((21)(\lambda(21)))$

1,2 ... De Bruijn indices

$\longrightarrow$ closed terms



Lambda terms

Let $\mathcal{V}$ be a countable set of variables. Lambda terms are defined by the following grammar:

$$\mathcal{T} ::= \mathcal{V} \mid \lambda \mathcal{V}.\mathcal{T} \mid \mathcal{T}\ \mathcal{T}$$

Example:

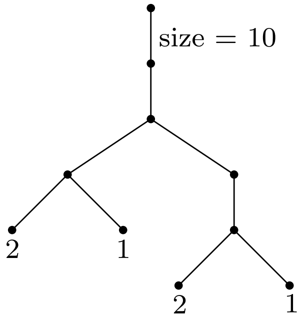
$\lambda x.\lambda y.((xy)(\lambda z.(yz)))$

De Bruijn notation:
$\lambda\lambda((21)(\lambda(21)))$

1,2 ... De Bruijn indices

$\longrightarrow$ closed terms

size = # vertices



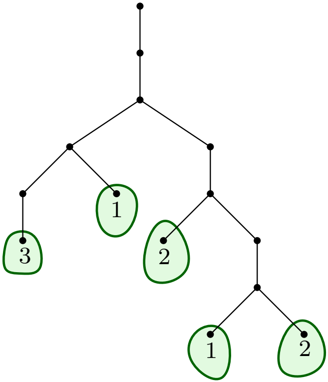
Restricted lambda terms

Bodini, Gardy, Gittenberger and Gołębiewski studied:

Restricted lambda terms

Bodini, Gardy, Gittenberger and Gołębiewski studied:

1. Lambda terms with bounded De Bruijn indices

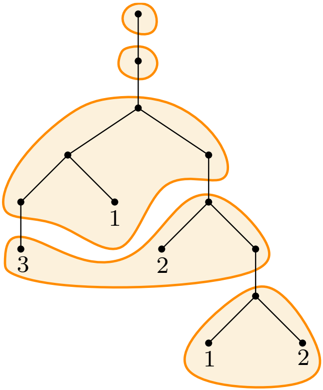


Restricted lambda terms

Bodini, Gardy, Gittenberger and Gołębiewski studied:

1. Lambda terms with bounded De Bruijn indices

2. Lambda terms with bounded number of De Bruijn levels



Total number of variables - Gittenberger, L.; 2019

Let X_n be the total number of variables in ...

...lambda terms where all De Bruijn indices are at most k.

Then $X_n \sim \mathcal{N}\left(\mu n, \sigma^2 n\right)$, with

$$\mu \sim \frac{\sqrt{k}}{2\sqrt{k}+1} \qquad \text{and} \qquad \sigma^2 \sim \frac{\sqrt{k}}{2(2\sqrt{k}+1)^2}.$$

Total number of variables - Gittenberger, L.; 2019

Let X_n be the total number of variables in ...

...lambda terms where all De Bruijn indices are at most k.

Then $X_n \sim \mathcal{N}\left(\mu n, \sigma^2 n\right)$, with

$$\mu \sim \frac{\sqrt{k}}{2\sqrt{k}+1} \qquad \text{and} \qquad \sigma^2 \sim \frac{\sqrt{k}}{2(2\sqrt{k}+1)^2}.$$

...lambda terms with at most $k+1$ De Bruijn levels.

If $B''(1) + B'(1) - B'(1)^2 \neq 0$, where $B(u) = \frac{\rho_k(u)}{\rho_k}$, with ρ_k denoting the dominant singularity, then

$$X_n \sim \mathcal{N}\left(\mu n, \sigma^2 n\right)$$

where $\mu = B'(1)$ and $\sigma^2 = B''(1) + B'(1) - B'(1)^2$.

Lambda terms with at most $k+1$ De Bruijn levels

$$N_i = u_i^2 - u_i + i, \quad \text{for all } i \geq 0,$$

where

$$u_0 = 0, \quad u_{i+1} = u_i^2 + i + 1 \quad \text{for } i \geq 0.$$

j	N_j	u_j
1	1	1
2	8	3
3	135	12
4	21760	148
5	479982377	21909
⋮	⋮	⋮

Lambda terms with at most $k+1$ De Bruijn levels

$$N_i = u_i^2 - u_i + i, \quad \text{for all } i \geq 0,$$

where

$$u_0 = 0, \quad u_{i+1} = u_i^2 + i + 1 \quad \text{for } i \geq 0.$$

j	N_j	u_j
1	1	1
2	8	3
3	135	12
4	21760	148
5	479982377	21909
⋮	⋮	⋮

Theorem (Bodini, Gardy, Gittenberger, Gołębiewski; 2018)

Let $H_k(z)$ be the generating function of lambda terms with at most $k+1$ De Bruijn levels.

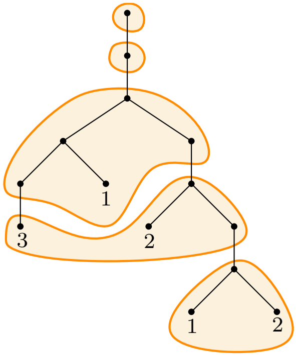
- *If $k \in (N_j, N_{j+1})$, then the dominant singularity ρ_k of $H_k(z)$ is of type $1/2$ and*
$$[z^n]H_k(z) \sim h_k n^{-3/2} \rho_k^{-n}.$$
- *If $k = N_j$, then the dominant singularity ρ_k of $H_k(z)$ is of type $1/4$ and*
$$[z^n]H_k(z) \sim h_k n^{-5/4} \rho_k^{-n}.$$

Lambda terms with at most $k+1$ De Bruijn levels

$$H_k(z) = \frac{1-\overbrace{\sqrt{1-2z+2z\sqrt{\ldots 1-4(k-1)z^2-2z+2z\sqrt{1-4kz^2}}}}^{k+1 \text{ nested square roots}}}{2z}$$

Lambda terms with at most $k+1$ De Bruijn levels

$$H_k(z) = \frac{\overbrace{1-\sqrt{1-2z+2z\sqrt{\ldots 1-4(k-1)z^2-2z+2z\sqrt{1-4kz^2}}}}^{k+1 \text{ nested square roots}}}{2z}$$



Lambda terms with at most $k+1$ De Bruijn levels

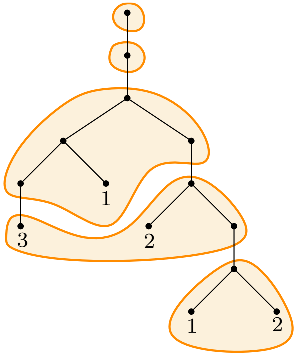
$$H_k(z) = \frac{1-\overbrace{\sqrt{1-2z+2z\sqrt{\ldots 1-4(k-1)z^2-2z+2z\sqrt{1-4kz^2}}}}^{k+1 \text{ nested square roots}}}{2z}$$

- $\boldsymbol{k \in (N_j, N_{j+1})}$:

  ρ_k comes from the $(j+1)$-th radicand $\longrightarrow$ type $\frac{1}{2}$

- $\boldsymbol{k = N_j}$:

  the j-th and $(j+1)$-th vanish both at ρ_k $\longrightarrow$ type $\frac{1}{4}$



Lambda terms with at most $k+1$ De Bruijn levels

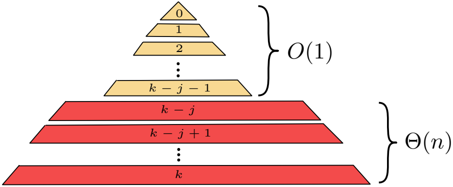
Gittenberger, L.; 2019:

Let X_n be the number of leaves at level l.

For $\boldsymbol{k \in (N_j, N_{j+1})}$

- $l < k - j : \mathbb{E}X_n \sim C_{k,l}$
- $l \geq k - j : \mathbb{E}X_n \sim \tilde{C}_{k,l} \cdot n$

with constants $C_{k,l}$ and $\tilde{C}_{k,l}$.



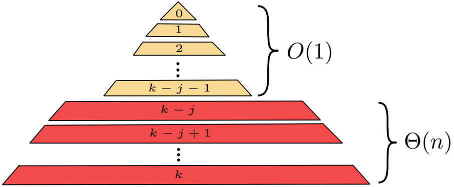
Lambda terms with at most $k+1$ De Bruijn levels

Gittenberger, L.; 2019:

Let X_n be the number of leaves at level l.

For $k \in (N_j, N_{j+1})$

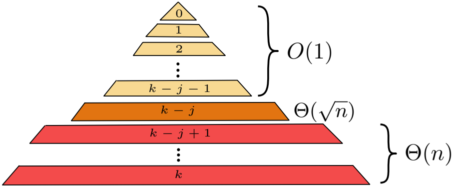
- $l < k-j : \mathbb{E}X_n \sim C_{k,l}$
- $l \geq k-j : \mathbb{E}X_n \sim \tilde{C}_{k,l} \cdot n$



with constants $C_{k,l}$ and $\tilde{C}_{k,l}$.

For $k = N_j$

- $l < k-j : \mathbb{E}X_n \sim D_{k,l}$
- $l = k-j : \mathbb{E}X_n \sim \hat{D}_{k,l} \cdot \sqrt{n}$
- $l > k-j : \mathbb{E}X_n \sim \tilde{D}_{k,l} \cdot n$



with constants $D_{k,l}$, $\hat{D}_{k,l}$ and $\tilde{D}_{k,l}$.

Lambda terms with at most $k+1$ De Bruijn levels

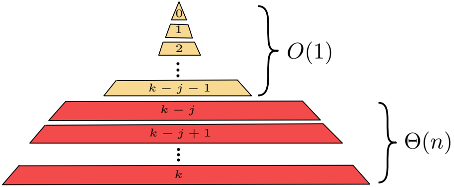
Gittenberger, L.; 2019:

Let X_n be the number of leaves at level l.

For $\boldsymbol{k \in (N_j, N_{j+1})}$

- $l < k-j : \mathbb{E}X_n \sim C_{k,l}$
- $l \geq k-j : \mathbb{E}X_n \sim \tilde{C}_{k,l} \cdot n$

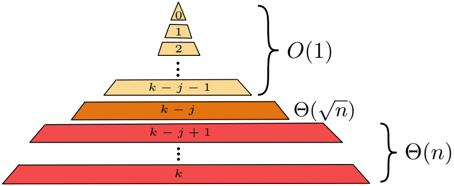
with constants $C_{k,l}$ and $\tilde{C}_{k,l}$.



For $\boldsymbol{k = N_j}$

- $l < k-j : \mathbb{E}X_n \sim D_{k,l}$
- $l = k-j : \mathbb{E}X_n \sim \hat{D}_{k,l} \cdot \sqrt{n}$
- $l > k-j : \mathbb{E}X_n \sim \tilde{D}_{k,l} \cdot n$

with constants $D_{k,l}$, $\hat{D}_{k,l}$ and $\tilde{D}_{k,l}$.



Lambda terms where all De Bruijn indices are at most k

$$G_k(z) = \frac{1 - \overbrace{\sqrt{1-4\cdot 0\cdot z^2-2z+2z\sqrt{\ldots 1-4(k-1)z^2-2z+2z^2+2z\sqrt{(1-z)^2-4kz^2}}}}^{k+1\text{ nested square roots}}}{2z}$$

Lambda terms where all De Bruijn indices are at most k

$$G_k(z) = \frac{1-\overbrace{\sqrt{1-4\cdot 0\cdot z^2-2z+2z\sqrt{\ldots 1-4(k-1)z^2-2z+2z^2+2z\sqrt{(1-z)^2-4kz^2}}}}^{k+1 \text{ nested square roots}}}{2z}$$

Theorem (Bodini, Gardy, Gittenberger, Gołębiewski; 2018)

Let $G_k(z)$ be the generating function of lambda terms where all De Bruijn indices are at most k. Then

- *the dominant singularity ρ_k of $G_k(z)$ comes from the innermost radicand* $\longrightarrow \rho_k = \frac{1}{2\sqrt{k}+1}$

Lambda terms where all De Bruijn indices are at most k

$$G_k(z) = \frac{1-\overbrace{\sqrt{1-4\cdot 0\cdot z^2-2z+2z\sqrt{\ldots 1-4(k-1)z^2-2z+2z^2+2z\sqrt{(1-z)^2-4kz^2}}}}^{k+1 \text{ nested square roots}}}{2z}$$

Theorem (Bodini, Gardy, Gittenberger, Gołębiewski; 2018)

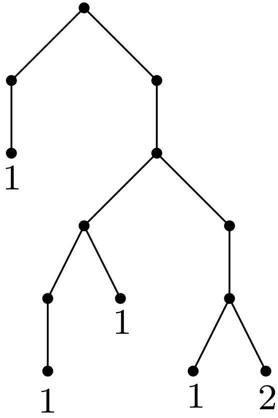
Let $G_k(z)$ be the generating function of lambda terms where all De Bruijn indices are at most k. Then

- *the dominant singularity ρ_k of $G_k(z)$ comes from the innermost radicand $\longrightarrow \rho_k = \frac{1}{2\sqrt{k}+1}$*
- *the asymptotics of the coefficients of $G_k(z)$ read as*

$$[z^n]G_k(z) \sim \frac{C_1^k k^{1/4}}{2\sqrt{\pi\rho_k}} n^{-3/2} \rho_k^{-n},$$

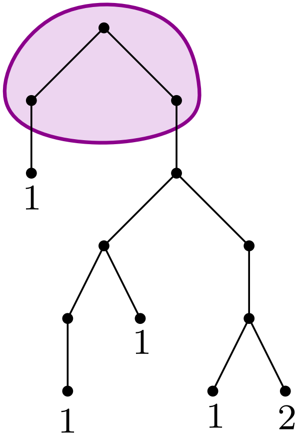
where $C_i^k := \prod_{s=i}^{k} \frac{1}{\sqrt{c_s}}$ with $c_1 = 5$ and $c_i = 4i - 5 + \sqrt{c_{i-1}}$.

Lambda terms where all De Bruijn indices are at most k



Lambda terms where all De Bruijn indices are at most k

$G_k(z) = B(z,$



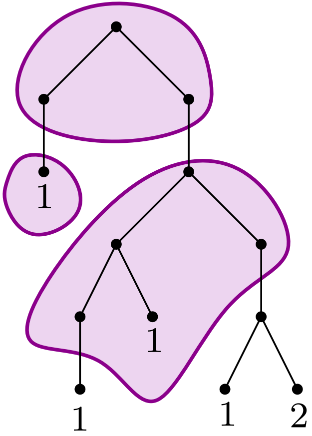
BGF of binary trees:

$B(z,u) = \frac{1-\sqrt{1-4uz^2}}{2z}$

$z \ldots$ size, $u \ldots$ leaves

Lambda terms where all De Bruijn indices are at most k

$G_k(z) = B(z, B(z,$



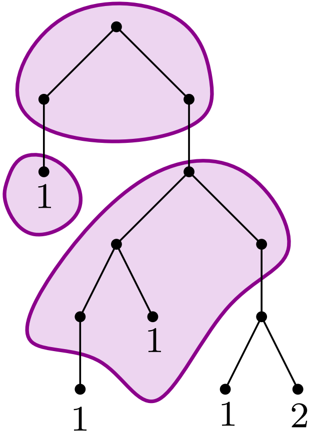
BGF of binary trees:

$B(z,u) = \frac{1-\sqrt{1-4uz^2}}{2z}$

$z \dots$ size, $u \dots$ leaves

Lambda terms where all De Bruijn indices are at most k

$G_k(z) = B(z, B(z, 1 + B(z,$



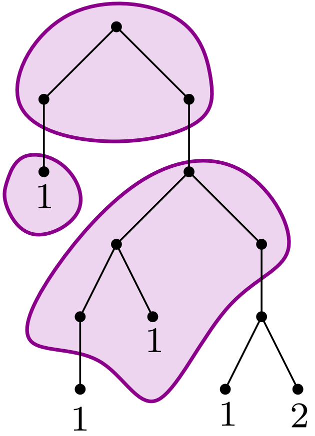
BGF of binary trees:

$B(z,u) = \frac{1-\sqrt{1-4uz^2}}{2z}$

$z \ldots$ size, $u \ldots$ leaves

Lambda terms where all De Bruijn indices are at most k

$$G_k(z) = B(z, B(z, 1 + B(z, 2 + \ldots + B(z, k-1 + B(z, k + B(z, \ldots$$



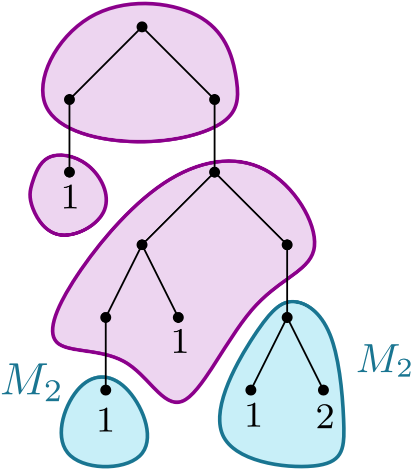
BGF of binary trees:

$B(z,u) = \frac{1-\sqrt{1-4uz^2}}{2z}$

$z \ldots$ size, $u \ldots$ leaves

Lambda terms where all De Bruijn indices are at most k

$$G_k(z) = B(z, B(z, 1 + B(z, 2 + \ldots + B(z, k-1 + M_k(z))))))$$



BGF of binary trees:

$B(z,u) = \frac{1-\sqrt{1-4uz^2}}{2z}$

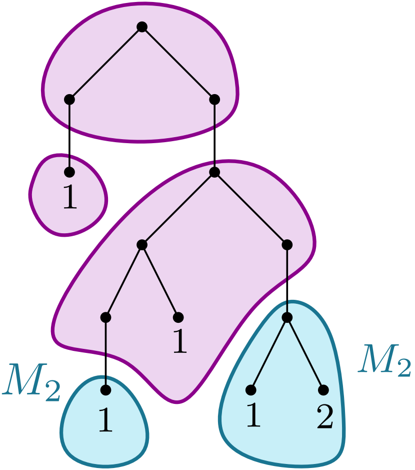
$z \ldots$ size, $u \ldots$ leaves

GF of k-colored Motzkin trees:

$M_k(z) = \frac{1-z-\sqrt{(1-z)^2-4kz^2}}{2z}$

Lambda terms where all De Bruijn indices are at most k

$$G_k(z) = B(z, B(z, 1 + B(z, 2 + \ldots + B(z, k-1 + M_k(z))))))$$



BGF of binary trees:

$B(z,u) = \frac{1-\sqrt{1-4uz^2}}{2z}$

$z \ldots$ size, $u \ldots$ leaves

GF of k-colored Motzkin trees:

$M_k(z) = \frac{1-z-\sqrt{(1-z)^2-4kz^2}}{2z}$

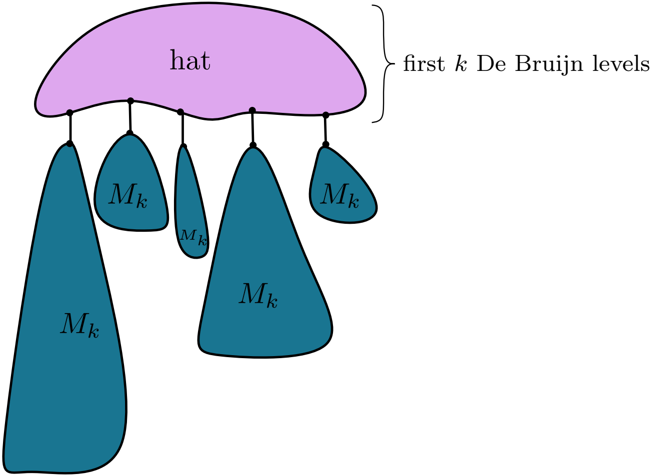
$\longrightarrow B(z, k + B(z,k)) = B(z,k)$

Lambda terms where all De Bruijn indices are at most k

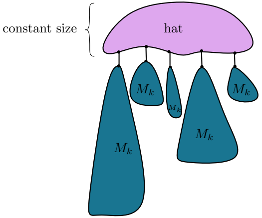
$$= k\text{-indexed lambda terms}$$

Lambda terms where all De Bruijn indices are at most k

$$= k\text{-indexed lambda terms}$$



Size of the hat



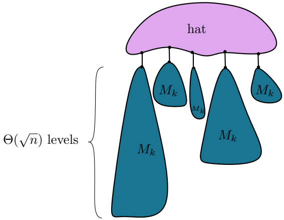
Theorem (Grygiel, L.; 2019$^+$)

For a fixed $k \geq 1$ let χ_k be the size of the hat of a k-indexed lambda term. Then

$$\mathbb{E}_{\mathcal{G}_{k,n}}(\chi_k) = k + 4(k + \sqrt{k} - 1) \sum_{i=1}^{k} \frac{C_1^i}{\sqrt{c_i}} + \sum_{i=0}^{k-2} \left(1 + 2\sqrt{k} + 4i - \sqrt{c_{k-i-1}}\right) \sum_{j=k-i}^{k} \frac{C_{k-i}^j}{\sqrt{c_j}} + o(1),$$

where $C_i^k := \prod_{s=i}^{k} \frac{1}{\sqrt{c_s}}$ *with* $c_1 = 5$ *and* $c_i = 4i - 5 + \sqrt{c_{i-1}}$.

Number of De Bruijn levels

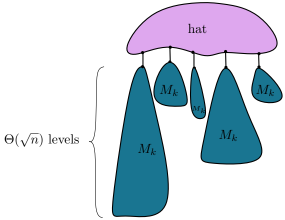


Theorem (Grygiel, L.; 2019$^+$)

The average number of De Bruijn levels of a k-colored Motzkin tree of size n is asymptotically equal to

$$\sqrt{\frac{\pi n}{2k + \sqrt{k}}}.$$

Number of De Bruijn levels



Theorem (Grygiel, L.; 2019^{+})

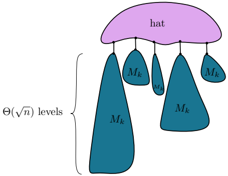
The average number of De Bruijn levels of a k-colored Motzkin tree of size n is asymptotically equal to

$$\sqrt{\frac{\pi n}{2k + \sqrt{k}}}.$$

Corollary

For every $k \geq 1$ the average number of De Bruijn levels in k-indexed lambda terms of size n is $\Theta(\sqrt{n})$.

Proof



Lemma (Drmota, De Mier, Noy; 2014)

Suppose that:

- *$F(z,t)$ is an analytic function at $(z,t) = (0,0)$*
- *$T(z) = F(z, T(z))$ has a "nice" solution $T(z)$*
- *$T(z)$ has a square-root singularity at $z = z_0$*
- *Let $T^{[k]}(z) = F(z, T^{[k-1]}(z))$.*
- *Let H_n be a random variable that is defined by*

$$\mathbb{P}\{H_n \le k\} = \frac{[z^n]T^{[k]}(z)}{[z^n]T(z)}.$$

Then

$$\mathbb{E}H_n \sim \sqrt{\frac{2\pi n}{z_0 F_z(z_0,t_0)F_{tt}(z_0,t_0)}}.$$

Unary profile

Theorem (Grygiel, L.; 2019$^+$)

Let $\kappa > 0$ be a fixed real number. The expected number of variables in the De Bruijn level $\lfloor \kappa\sqrt{n} \rfloor$ in k-indexed lambda terms of size n is asymptotically equal to

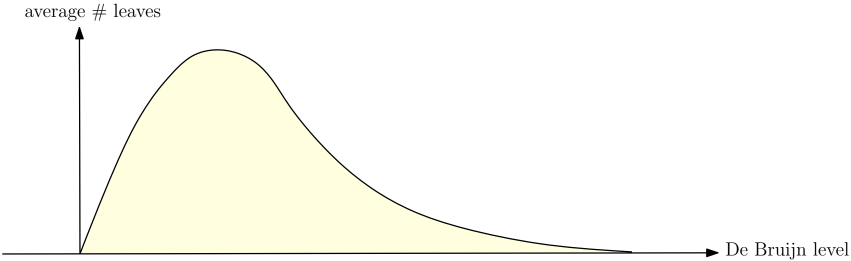
$$2\kappa \exp\left(-\kappa^2(2k + \sqrt{k})\right)\sqrt{n}.$$

Unary profile

Theorem (Grygiel, L.; 2019^+)

Let $\kappa > 0$ be a fixed real number. The expected number of variables in the De Bruijn level $\lfloor \kappa\sqrt{n} \rfloor$ in k-indexed lambda terms of size n is asymptotically equal to

$$2\kappa \exp\left(-\kappa^2(2k+\sqrt{k})\right)\sqrt{n}.$$

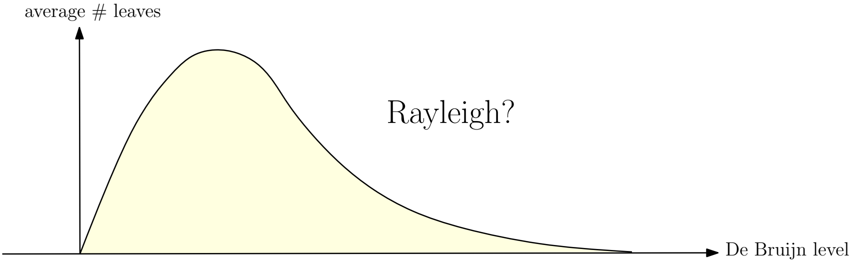


Unary profile

Theorem (Grygiel, L.; 2019^{+})

Let $\kappa > 0$ be a fixed real number. The expected number of variables in the De Bruijn level $\lfloor \kappa\sqrt{n} \rfloor$ in k-indexed lambda terms of size n is asymptotically equal to

$$2\kappa \exp\left(-\kappa^2(2k + \sqrt{k})\right)\sqrt{n}.$$



Idea of proof

Tool: bivariate generating functions

Idea of proof

Tool: bivariate generating functions $\longrightarrow$ correct marking:

Idea of proof

Tool: bivariate generating functions $\longrightarrow$ correct marking:

- size of the hat

$$G_k(z,\boldsymbol{u}) = B(z\boldsymbol{u}, B(z\boldsymbol{u}, 1+B(z\boldsymbol{u}, 2+\ldots+B(z\boldsymbol{u}, k-1+M_k(z))))))$$

Idea of proof

Tool: bivariate generating functions $\longrightarrow$ correct marking:

- size of the hat

$$G_k(z, \boldsymbol{u}) = B(z\boldsymbol{u}, B(z\boldsymbol{u}, 1{+}B(z\boldsymbol{u}, 2{+}\ldots{+}B(z\boldsymbol{u}, k{-}1{+}M_k(z))))))$$

- leaves in De Bruijn level $\lfloor \kappa\sqrt{n} \rfloor$

$$G_k(z, \boldsymbol{u}) = B\Big(z, B\Big(z, 1 + B\Big(z, 2 + \ldots$$
$$+ \underbrace{B(z, k + B(z, k + B(\ldots B}_{\kappa\sqrt{n}\text{ occurrences of } B}(z, k + B(z, k\boldsymbol{u} + M_k(z)))))) \ldots \Big)\Big)\Big)$$

Idea of proof

Tool: bivariate generating functions $\longrightarrow$ correct marking:

- size of the hat

$$G_k(z,u) = B(zu, B(zu, 1+B(zu, 2+\ldots+B(zu, k-1+M_k(z))))))$$

$$\longrightarrow \left.\frac{\partial G_k(z,u)}{\partial u}\right|_{u=1} = A_k - B_k\sqrt{1-\frac{z}{\rho_k}} + \mathcal{O}\left(\left|1-\frac{z}{\rho_k}\right|\right)$$

- leaves in De Bruijn level $\lfloor \kappa\sqrt{n} \rfloor$

$$G_k(z,u) = B\Big(z, B\Big(z, 1 + B\Big(z, 2 + \ldots$$
$$+ \underbrace{B(z, k + B(z, k + B(\ldots B}_{\kappa\sqrt{n} \text{ occurrences of } B}(z, k + B(z, ku + M_k(z)))))) \ldots \Big)\Big)\Big)$$

Idea of proof

Tool: bivariate generating functions $\longrightarrow$ correct marking:

- size of the hat

$$G_k(z, \boldsymbol{u}) = B(z\boldsymbol{u}, B(z\boldsymbol{u}, 1+B(z\boldsymbol{u}, 2+\ldots+B(z\boldsymbol{u}, k-1+M_k(z))))))$$

$$\longrightarrow \left.\frac{\partial G_k(z,u)}{\partial u}\right|_{u=1} = A_k - B_k\sqrt{1-\frac{z}{\rho_k}} + \mathcal{O}\left(\left|1-\frac{z}{\rho_k}\right|\right)$$

- leaves in De Bruijn level $\lfloor \kappa\sqrt{n} \rfloor$

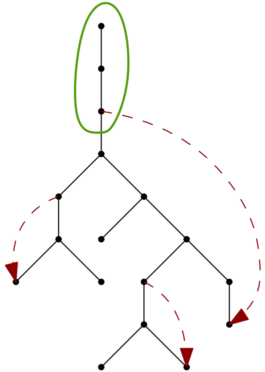
$$G_k(z, \boldsymbol{u}) = B\Big(z, B\Big(z, 1 + B\Big(z, 2 + \ldots$$

$$+ \underbrace{B(z, k + B(z, k + B(\ldots B}_{\kappa\sqrt{n}\text{ occurrences of } B}(z, k + B(z, k\boldsymbol{u} + M_k(z)))))) \ldots \Big)\Big)\Big)$$

$$\longrightarrow [z^n]\left.\frac{\partial G_k(z,u)}{\partial u}\right|_{u=1} = \frac{1}{2\pi i}\int_\gamma \frac{z^{k+\kappa\sqrt{n}-n-1}}{\prod_{j=1}^{k+\kappa\sqrt{n}}\sqrt{R_{k,j}(z,1)}}dz$$

Some open problems

- Enumeration and structure of these lambda terms for k depending on n
- Average number of binding lambdas
- Average number of starting lambdas



Some open problems

- Enumeration and structure of these lambda terms for k depending on n
- Average number of binding lambdas
- Average number of starting lambdas